Accident Analysis and Prevention 58 (2013) 106–114

Contents lists available at SciVerse ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

The effect of Operation 24 Hours on reducing collision in the City of Edmonton

Siana Halim[a,*], Heming Jiang[b,1]

[a] Industrial Engineering Department, Petra Christian University, Jl. Siwalankerto 21-131, Surabaya, Indonesia
[b] City of Edmonton Office of Traffic Safety, 9304 – 41 Ave, Edmonton T6E 6G8, Alberta, Canada

ARTICLE INFO

Article history:
Received 21 February 2013
Received in revised form 21 April 2013
Accepted 25 April 2013

Keywords:
Negative binomial time series
Segmented regression
Change point

ABSTRACT

In the City of Edmonton, in order to reduce the prevalence of collisions, the Operation 24 Hours program (OPS24) was developed by using existing police and transportation services resources. The program uses traditional manned police speed enforcement method, which are supplemented by traffic safety messages displayed on permanent and mobile dynamic messaging signs (DMS). In this paper, collision data analysis was performed by looking at the daily number of collisions from 2008 to 2011 that covers 28 Operation 24 Hours (OPS24) events. The objective of the collision data analysis is to analyze if there is a reduction in collision frequencies after OPS24 was held and examined how long the collision reduction effect last. Weather factors such as temperature, thickness of snow, and wind gust have been considered by many as a great influence on collision occurrences, especially in a city with long and cold winter such as Edmonton. Therefore, collision modeling was performed by considering these external weather factors. To analyze the linear and periodic trend of different collision types (injury, fatal, and property damage only (PDO)) and examine the influence of weather factors on collisions, negative binomial time series model that accounts for seasonality and weather factors was used to model daily collision data. The modeling also considered collision proportion to account for missing traffic volume data; the Gaussian time series model that accounts for seasonality and weather factors was used to model collision proportion. To estimate the collision trend and test for changes in collision levels before/after OPS24, interrupted time series model with segmented regression was used. While for estimating how long the effect of the OPS24 last, change point method was applied.

1. Introduction

Many researchers have considered weather factors as a great influence on collision occurrences. Poor weather-related driving conditions are associated with 7000 fatalities, 800,000 injuries, and more than 1.5 million vehicular collisions annually in the United States (National Research Council, 2004). Adverse weather is present in 28% of total collisions and nearly 20% of highway fatalities (Weather and Highways, 2004). Analysts estimate the economic toll of weather-related collisions at $42 billion (Lombardo, 2000). In Canada, weather-related collisions have been estimated to cost Canadians an average of approximately CDN$ 1 billion per year (Andrey et al., 2001; Usman et al., 2011). Understanding the effects of adverse weather on motor vehicle collisions matters because experts have identified a number of communications and engineering innovations (largely technologies to collect and communicate real time road condition information, such as sensors and dynamic message signs) that could significantly reduce the collision and injury rates, but at a potentially substantial cost (National Research Council, 2004; Weather and Highways, 2004).

It is a known fact that collisions can be influenced by external environmental variables such as temperature, snowfall level, and wind gust. Previous studies have associated precipitation with markedly increased collisions rates (Eisenberg, 2004; Brodsky and Hakkert, 1988; Andrey and Yagar, 1993; Fridstrom et al., 1995). Recent work also shows that the risk posed by precipitation rises dramatically with the time since last precipitation (Eisenberg, 2004). There has not been any study on the effect of snowfall.

Number of collisions is not inevitably higher in snowy weather than in dry weather. On the one hand, snow makes driving more dangerous; by reducing tire adherence and impairing visibility. On the other hand, experienced drivers typically drive more slowly and carefully in snowy weather, and many people avoid or postpone unnecessary travel. Perhaps as a reflection of these offsetting factors, the handful of published studies addressing the collision consequences of snow has produced some conflicting results. The

* Corresponding author. Tel.: +62 31 298 3425.
E-mail addresses: halim@peter.petra.ac.id, sianahalim@yahoo.com (S. Halim), heming.jiang@edmonton.ca (H. Jiang).
[1] Tel.: +1 780 495 9906.

0001-4575/$ – see front matter
http://dx.doi.org/10.1016/j.aap.2013.04.031

Table 1
2008–2011 OPS24 schedules.

2008	2009	2010	2011
	Thursday, January 29	Wednesday, January 20	Wednesday, January 19
	Tuesday, March 10	Friday, February 05	Friday, February 04
	Tuesday, April 07	Monday, March 29	Monday, March 28
	Tuesday, May 12	Tuesday, April 13	Thursday, April 21
	Tuesday, June 09	Monday, June 28	Thursday, May 19
	Friday, June 26		Tuesday, June 07
Monday, September 15	Tuesday, September 29	Monday, September 13	
Tuesday, October 21	Friday, October 16	Monday, November 15	Tuesday, November 15
Tuesday, December 16	Tuesday, December 15	Tuesday, December 14	Friday, December 16

weight of the evidence suggests that less severe collisions (e.g., those producing only property damage) increase during snowfall, while more severe collisions (those resulting in major injuries or fatalities) decrease.

Significantly increased collision rates have been documented in snowy months in Canada (Andreescu and Frost, 1998), on snowy days in the United Kingdom (Perry et al., 1991), and during snowstorms in Iowa, United States (Knapp et al., 2000). Perry et al. (1991) found increased rates of collisions involving injuries and fatalities on snowy days in the United Kingdom, but Brown and Baass (1997) noted fewer collisions involving injuries in the winter months in Canada, as did Fridstrom et al. (1995) in snowy months in Denmark and Finland. Eisenberg (2004) found decreased rates of fatal collisions on snowy days in the United States, a finding echoed in analysis of winter months in Canada (Brown and Baass, 1997) and snowy months in Scandinavia. To date, only two previous studies have examined the effects of the first snowfall of the season. Defining first snowfall as the first snow in a month following a month without snow, Fridstrom and Ingebrigtsen (1991) found significant increases in both injury and fatal collisions in Norway. Subsequently, however, research by Fridstrom et al. (1995) produced mixed findings: Injury collisions rose significantly during the winter's first month with snow (compared with other months with snow) in Denmark but not in either Finland or Norway. Fatality rates were no different in the first snowy month than in other snowy months.

1.1. *The Operation 24 Hours*

In the City of Edmonton, in order to reduce the prevalence of speed related collisions, the Operation 24 Hours program (OPS24) was developed by using existing police and transportation services resources. The program uses traditional police speed enforcement methods such as manned enforcement, which are supplemented by traffic safety messages displayed on permanent and mobile dynamic messaging signs (DMS). The program utilizes an alternating two-message approach "Big Ticket Event – Don't Speed". The messages were run consecutively for four days prior to the operation day and also during the operation day (the fifth day). During the OPS24, the entire City of Edmonton became an enforcement zone. Twenty-eight OPS24 events have been conducted during 2008–2011 (see Table 1).

The objectives of this study are to examine the impact of OPS24 on collisions in general as well as on different collision types (injury, fatal, and property damage only (PDO)). The collision analysis did not limit to speed related collisions since the Edmonton Police Service collected speed information of vehicles involved in a collision only starting from 2011 and only if the police visited the collision scene. The collected speed information is also only categorical: unsafe speed, no unsafe speed, unknown, and not applicable. General collision trend and the influence of weather factors such as temperature, thickness of snow, and wind gust were also investigated. In addition, we aim to find how long the collision reduction lasts. The impact of OPS24 on vehicle speed distribution is beyond the scope of this paper and can be found in Halim et al. (2012).

2. Data

This study used 3 years and 4 months of daily collision data ranging from September 2008 to December 2011. The collision data used in this analysis was obtained from the City of Edmonton's Motor Vehicle Collision Information System (MVCIS), a database of motor vehicle collisions that occur on public roads in the City of Edmonton. The information in the database is collected from the provincial Collision Report Form, which is completed by members of the Edmonton Police Service (EPS) either on paper at the scene of the collision or electronically at the front counter of a divisional or community police station. The collision reports received from the EPS were reviewed and corrected if necessary (in consultation with the EPS) before entered into MVCIS.

The total number of collisions was 109,459 during this period. The daily collision dataset contains collision frequencies, collision types (injury, fatal, and PDO), collision locations, roadway portions, and collision causes. We combined fatal and injury collisions and named it 'severe collisions' which contributes 14.6% to the total collisions. The statistical summary of the data is presented in Table 2.

2.1. *Weather data*

The weather data used in this analysis includes 3 years and 4 months of daily weather data of the City of Edmonton from Environment Canada with dates ranging from September 2008 to December 2011. The daily weather data contains variables such as daily maximum, minimum and mean temperatures in degree Celsius, the thickness of snow on the ground (in cm), and the speed of maximum wind gust during the day (in km/h). The statistical summary of the weather data is shown in Table 3.

Both MVCIS and Environment Canada provide reliable data for purpose of this study. The completeness and accuracy of data are significantly high, for example, no unknown information about the

Table 2
Summary of collision data.

	Minimum	Maximum	Mean	Std deviation
Number of collisions per day	13	305	75.13	35.24
Severe collisions per day	1	36	10.98	4.69
PDO collisions per day	12	286	64.08	32.63

Table 3
Summary of weather data.

	Minimum	Maximum	Mean	Std deviation
Min temperature (°C)	−46.1	14.3	−5.49	11.99
Max temperature (°C)	−33.2	34.0	8.46	13.21
Mean temperature (°C)	−39.7	22.8	1.48	12.34
Ground snow (depth, cm)	0	52	7.52	12.31
Speed of max gust (km/h)	4	95	37.88	10.95

collision type and the percentage of unknown collision causes is only less than 1%.

Fig. 1 shows seasonal variation of collisions and their relationship to weather. The collisions in autumn and winter are higher than in spring and summer, it had a periodic trend. We can also see that when the mean temperature was low, the collisions were high and vice versa. Fig. 1 also shows that when the snow on the ground was low, the collisions were low. There was no clear relationship between the speed of maximum wind gust and the collisions.

3. Methods

We divided collision and weather data analysis into three sections. First, we modeled the collision trends and analyzed the effects of environmental variables such as temperature, thickness of snow, and speed of maximum wind gust on collision frequencies and collision proportions. For the collision frequencies, we used negative binomial time series model. Since traffic volume data were not available, we also looked at the proportions of daily injury and fatal or PDO collisions. Those are expected to be independent of the traffic volume. For these collision proportions, we used the Gaussian time series model that was designed to handle continuous dependent variables.

Some models for collisions have been constructed based on time series analysis. Vanlaar et al. (2011) modeled the log transform of monthly collisions from 1994–2008 at 48 intersections using ARIMA (p, d, q) models. Quddus (2008) modeled traffic collisions in Great Britain using class of INAR models for time series. He used two sets of data to test his models, that is, annual traffic collisions in Great Britain from 1950 to 2005 and monthly traffic collision casualties within the London congestion-charging zone between January 1991 and October 2005. Another approach proposed by Castro et al. (2012) was based on partitioning of a single latent continuous variable into mutually exclusive intervals. By this way, Castro et al. (2012) reformulated daily count models as a special case of generalized ordered-response models. This model was applied to predict collision frequencies at urban intersections in Arlington, Texas. Models offered by Vanlaar et al. (2008) and Quddus (2008) are only for univariate data and the model fittings in these papers were applied only for short series or aggregated series, whereas Castro et al. (2012) introduced spatial and temporal dependencies through the latent continuous variables. A review and assessment of methodologies in the statistical analysis of collisions is given by Lord and Mannering (2010). The negative binomial model was also chosen over the time series analysis models since it is a multivariate model rather than a univariate one.

Moreover, the negative binomial time series is chosen since for time series consisting of counts, classical Gaussian models are inappropriate. Count data are non-negative, integer-valued and often overdispersed, i.e., the variance is larger than the mean. Poisson models are not robust for modeling count data, since they have over dispersion problems (Cameron and Trivedi, 1998). Therefore, many researchers used overdispersed Poisson and binomial regression model to handle this problem. As a natural extension of Poisson distribution, it is well known that negative binomial distribution is more flexible and allows for overdispersion (Davis and Wu, 2009). El-Basyouny and Kwon (2012) modeled the time and weather effects on collision frequencies using City of Edmonton data from 2000 to 2010 by using multivariate Poisson lognormal models. This model is more robust then the Poisson models. However, the Poisson-lognormal has limitations, i.e., the model estimation is more complex because the Poisson-lognormal distribution does not have a closed form. Therefore, many researchers used the Markov Chain Monte Carlo simulation for estimating the parameters of those models, since the classical estimation of those parameters cannot be calculated in a straightforward way. Moreover, the Poisson-lognormal can still be adversely affected by small sample sizes and low sample mean values (Miaou and Mallick, 2003). The negative binomial time series can be solved using classical estimation (Davis and Wu, 2009).

Secondly, we estimated the collision trend and tested for changes in collision levels before/after OPS24 by using interrupted time series with segmented regression (Perrin, 2009).

Finally, we measured the statistical properties, i.e., a sequence of collision means after OPS24 ended and estimated a point at which the statistical property of that sequence is changed. That point is called a change point (Killick et al., 2011). From the length of days between the ends of OPS24 to that change point, we can measure how long the behavior lasts in term of collisions reduction. Moreover, the change point detection models are usually used in the statistical process control for detecting a change in the distribution from the target in the process (Hawkins and Deng, 2010).

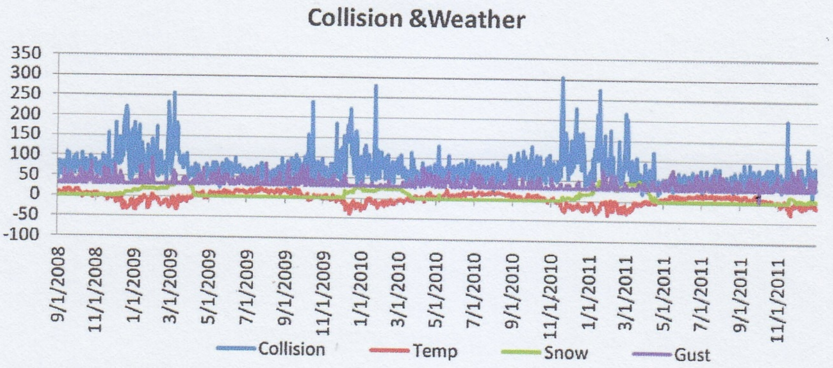

Fig. 1. The reported collision and weather in the Edmonton from September 2008 to December 2011.

The detection of a single change point can basically be regarded as a hypothesis test; with null hypothesis being there is no change point versus the alternative hypothesis being there exists a change point. Some models that assume normality in the observations are given by Hawkins et al. (2003), which detected the change in mean, and Hawkins and Zamba (2005), which detected change in mean and variance. Here, we applied the Hawkins and Deng (2010) nonparametric change point detection model. The non-parametric approach was chosen because we do not need to assume any distribution of the dataset as in a parametric model.

3.1. Negative binomial model (Davis and Wu, 2009)

Let Y_t be a time series of counts and X_t be an observed l-dimensional covariate, $X_t = \left(1, \frac{t}{N}, \cos\left(\frac{2\pi t}{360}\right), \sin\left(\frac{2\pi t}{360}\right), Temp_t, Snow_t, Gust_t\right)$, t is the discrete time, s is the seasonal term, and the link function is:

$$\log\left\{\frac{r(1-p_t)}{p_t}\right\} = x_t^T\beta + \alpha_t \tag{1}$$

where $\{\alpha_t\}$ latent process with Gaussian innovations. Here β is the vector of regression coefficients. The used of harmonic functions in the covariate X_t is for capturing the seasonal effects in the data (Davis et al., 2000).

The conditional density function of Y_t is

$$p(Y_t = y_t|\alpha_t) = \begin{pmatrix} y_t + r - 1 \\ r - 1 \end{pmatrix} p_t^r(1-p_t)^{y_t} \tag{2}$$

For $y_t = 0, 1, \ldots$ Let the process $\{\varepsilon_t = e^{\alpha_t}\}$, which is a strictly stationary nonnegative time series when $\{\alpha_t\}$ is strictly stationary, center $\{\alpha_t\}$ such that $\{\varepsilon_t\}$ has mean one. The conditional mean of Y_t can be written in terms of ε_t,

$$E(Y_t|\alpha_t) = \frac{r(1-p_t)}{p_t} = r\exp\left(x_t^T\beta + \alpha_t\right) = r\exp(x_t^T\beta)\varepsilon_t\infty \tag{3}$$

with $E(\varepsilon_t) = 1$, we have $E(Y_t) = r\exp(x_t^T\beta)$ the form for a pure generalized linear model in the negative binomial case. When $\{\alpha_t\}$ is, in order to satisfy the condition of $E(e^{\alpha_t}) = 1$, it is required that $\alpha_t \sim N(-\sigma_\alpha^2/2, \sigma_\alpha^2)$.A standard negative binomial generalized linear model was fitted to the data. The estimate $\hat{r}$ was determined by the r value that yielded the smallest AIC. The $\{\alpha_t\}$ was estimated from the residuals of (3). When the $\{\alpha_t\}$ was a stationary Gaussian linear process with innovations Z_t which are independent identically distributed $N(0, \sigma^2)$ variables. Then the σ^2 was estimated using the method of moments.

$$var(Y_t) = \mu_t + \mu_t^2\frac{(r+1)\sigma_\varepsilon^2 + 1}{r} \tag{4}$$

and the estimate of $\hat{\sigma}_\varepsilon^2$ and $\hat{\rho}_\varepsilon(k)$ were obtained using an ordinary least square type estimators (Brännäs and Johansson, 1994)

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{\hat{r}+1}\left[\frac{\hat{r}\sum_n^{t=1}\hat{\mu}_t^2\left\{(Y_t - \hat{\mu}_t)^2 - \hat{\mu}_t\right\}}{\sum_n^{t=1}\hat{\mu}_t^4} - 1\right] \tag{5}$$

$$\hat{\rho}_\varepsilon(k) = \frac{\hat{\sigma}_\varepsilon^{-2}\sum_n^{t=k+1}\hat{\mu}_t\hat{\mu}_{t-k}(Y_t - \hat{\mu}_t)(Y_{t-k} - \hat{\mu}_{t-k})}{\sum_n^{t=k+1}\hat{\mu}_t^2\hat{\mu}_{t-k}^2} \tag{6}$$

The latent process models' parameters can be estimated using (6) and the standard methods as, Yule Walker Estimate, algorithm for estimating moving average parameters (see Box and Jenkins, 1970).To get the solution of this model for predicting the collisions in Edmonton, we used R-programming (CRAN, 2012).

3.2. Gaussian time series model

To account for missing traffic volume data, the modeling also considered collision proportions. Since the Negative Binomial time series model is designed for count dependent variables, it will not be appropriate to handle the daily injury and fatal or PDO collision proportions. Therefore, we changed the link function to Gaussian, which is appropriate to handle continuous dependent variables.

3.3. Interrupted time series with segmented regression model

Study of interrupted time series has been applied in many traffic conditions; Hamed et al. (1999) used a diffusion model with jumps to investigate the effect of Gulf crisis to traffic collisions in Jordan. Here, we used a segmented regression (Perrin, 2009) for estimating the trend and level of collisions before/after OPS24. This model is more classical than the diffusion model. However, it can be used to examine the impact of external variables, such as temperature and snow on the ground to the change of trend and level of collisions before/after OPS24. The diffusion model is appropriate for modeling a univariate variable. The segmented regression can be formulated as follow:

$$Y_t = b_0 + b_1T + b_2D + b_3P + b_4Temp_t + b_5Snow_t + e_t, \tag{7}$$

where Y_t is the proportion of injury and fatal collisions; T is time, D is a dummy variable for pre or post intervention, P is time after the intervention, e_t is the random variation at time t not explained by the model. $Temp_t$ is the mean of temperature at the observation time t, b_i, i = 0,...,5 are coefficients that should be estimated.

3.4. Change point detection model

Hawkins and Deng developed non-parametric approach of change point models, by assuming $Y_1, Y_2, \ldots, Y_\tau, Y_{\tau+1}, \ldots, Y_n$ as independent, continuous random variables with statistical distribution

$$\begin{aligned} &Y_i \sim F(y)\ \text{for}\ i = 1, 2, \ldots, \tau \\ &Y_i \sim F\left(y - \theta\right)\ \text{for}\ i = \tau+1, \ldots, n, \end{aligned} \tag{8}$$

where θ is a shift in location occurring after the change point τ and n is the number of daily speed data for a given location. In this case, Y_i is the number of collisions started after the OPS24.

In the non-parametric approach, the likelihood ratio test for the null hypothesis that; $H_0 : \mu_1 = \mu_2$ (the mean of before and after the change point is the same) is the two-sample t-statistic. It is defined as:

$$T_{k,n} = \frac{U_{k,n}}{\sqrt{k(n-k)(n+1)/3}}; \tag{9}$$

with $U_{k,n} = 2\sum_{i=1}^{k}R_i - k(n+1)$; $1 \le k \le n-1$ and R_i is the rank of $Y_i$within $Y_1, \ldots, Y_n$,In the null case, we are assuming constant variance, the $T_{k,n}$ follows a t-distribution with $n-2$ degrees of freedom. The τ is given by finding $T_{\max,n}$, the maximum of $\left|T_{k,n}\right|$ across all possible k values.

4. Results

The results of the collision data analysis are separated into three sections. First, we present the results from general collision trend and weather influence analysis using negative binomial time series model. In the second section, we present the collision proportion analysis using gaussian time series model. Finally, we discuss the impact of OPS24 on collisions using interrupted time series with segmented regression model.

Table 4
Parameters of negative binomial model.

Explanatory variables	All collisions Coeff.	Pr(>|t|)	Injury and fatal collisions Coeff	Pr(>|t|)	PDO Coeff	Pr(>|t|)
Intercep	4.3131	<2e-16*	2.4201	<2e-16*	8.473e-01	<2e-16*
Slope	−0.2159	2.39e-13*	−0.2252	1.55e-11*	−7.777e-04	0.8489
Cosinus	0.1709	9.22e-10*	0.1844	7.95e-09*	−8.535e-03	0.027+
Sinus	0.0448	0.0240+	0.0354	0.11713	3.286e-03	0.2361
Temperature	-0.018	<2e-16*	−0.0037	0.05635#	−1.623e-03	3.13e-11*
Snow	0.0054	2.13e-05*	0.0045	0.0021−	2.966e-05	0.8682
Gust	0.0022	0.0226+	0.0009	0.40756	1.265e-04	0.3461
Null deviance	1842.5		1331.9		3.7317	
Residual deviance	1238.5		1228.9		2.9317	
AIC	11179		6893.6		−3867.	

Signif. codes:
* <0.01.
\+ 0.01.
− <0.05.
\# <0.10.

4.1. General collision trend and weather influence analysis

Firstly, we fitted the collision data using a standard negative binomial generalized linear model and the $\hat{r}$ was determined by the r value that yielded the smallest AIC. AIC stands for 'Akaike Information Criterion', it is a measure of the relative goodness of fit of a statistical model. The AIC is grounded in the concept of information entropy, in effect offering a relative measure of the information lost when a given model is used to describe reality. It can be said to describe the tradeoff between bias and variance in model construction, or loosely speaking between accuracy and complexity of the model. AIC values provide a means for model selection. The smaller the AIC value, the better the model fit. The parameters from this model were summarized in Table 4.

From Table 4, we can see that for all daily collisions, there was a significant declining annual trend; the periodic trend followed a cosine function, meaning that collisions decreased from January to June, and increased to December; there was a significant inverse relationship with mean temperature; significant positive relationship with thickness of snow and wind gust. For daily injury and fatal collisions, there was also a significant declining annual trend; the periodic trend also follows a cosine function; non-significant inverse relationship with mean temperature; significant positive relationship with thickness of snow; and non-significant positive relationship with wind gust. PDO collisions did not have a trend; it also followed a cosine function. It had a significant inverse relationship with the mean temperature, so we can say that the PDO collisions increase during the wintertime. The relationship with the thickness of snow and wind gust was positive but not statistically significant.

Now, using the above parameters, we then calculated the latent process α_t using Eqs. (7) and (8). Since the latent process is a random series, we did a Monte Carlo simulation for the latent process by generating that process $B = 1000$ times. Further, we fitted the mean of the estimated series, i.e., $\hat{Y}_t$ into the data and constructed the confidence interval by taking the 2.5% and 97.5% of the ordered estimated series. The fitted values and the confidence intervals for daily, injury and fatal and PDO collisions are depicted in Fig. 2(a–c) respectively. From those figures we can see that the models followed the seasonality trend. Except for the Injury and Fatal collisions, the confidence interval of the fitted value covered the data well.

The fitted model measurements are given in five types, i.e., mean square error (MSE), root mean square error (RMSE), mean of absolute deviation (MAD), symmetric mean absolute percentage error (SMAPE) and Theil's U1. SMAPE is an accuracy measure based on percentage errors. SMAPE can be formulated as $\sum_{t=1}^{N}|Y_t - \hat{Y}_t|/(Y_t + \hat{Y}_t)$, where Y_t is the true value of collisions and $\hat{Y}_t$ is the estimated one at time t. The Theil's U1 statistic is bound between 0 and 1, with values closer to 0 indicating greater forecasting accuracy. The Theil's U statistics measure whether the model being used is providing valuable information.

The summary of those measurements for the mean of the estimated series is given in Table 5. In which we can see that the SMAPE of the injury and fatal collisions only reach 15.9%, while for the daily collisions and PDO collisions it can reach 13.5% and 14.2%, respectively. Since the Theil's U1 statistic is closer to 0, we can deduce that the models are providing valuable information. Among four measurements, i.e., MSE, RMSE, MAD, SMAPE, the smaller the value, the better the fitted model. However, MSE, RMSE, and MAD are not robust to the outliers in the data. SMAPE is more robust to the outliers than those other three measurements (Flores, 1986).

4.2. Collision proportion analysis

To analyze the daily proportion of injury and fatal collisions or PDO collisions to total collisions, we used the GLM with Gaussian density function as the link function since the dependent variable is continuous. As shown in Table 6, for daily PDO and injury and fatal collision proportions, there was a non-significant declining annual trend; fairly significant cosine periodic trend; significant negative relationship with mean temperature; and non-significant positive relationship with thickness of snow and wind gust. The fitted model

Table 5
The fitted model measurements.

Measurement	Negative binomial time series: Daily collision	Injury and fatal collision	PDO
MSE	824.6126	17.9288	689.3258
RMSE	28.7161	4.2343	26.2550
SMAPE	13.505	15.87	14.15
MAD	20.3815	3.3675	18.1344
Theil U1	0.1785	0.1910	0.1899

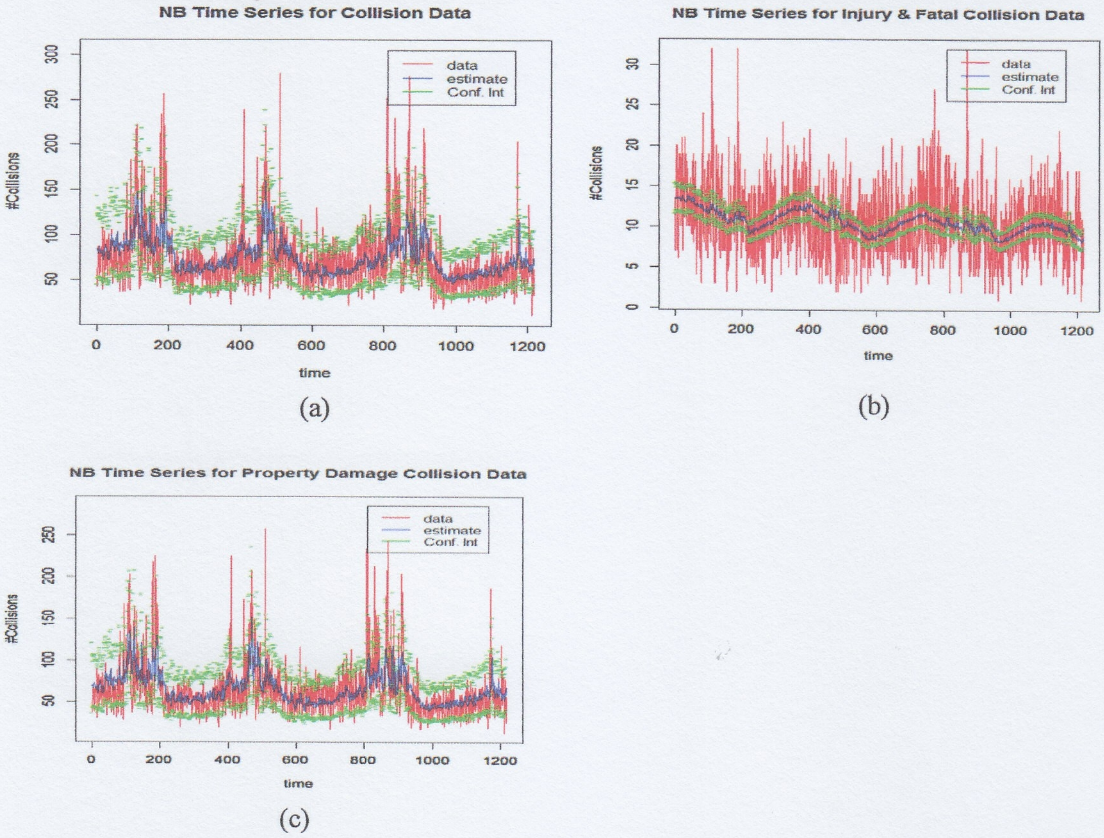

Fig. 2. Fitted model into the data for (a) daily collisions, (b) injury and fatal collisions, (c) PDO collisions.

to the actual proportion of the injury and fatal collision proportions and the PDO collision proportions are depicted in Fig. 3. In those figures we can see that the 95% confidence interval of both fitted models can cover the actual proportions.

In general, the PDO collision frequencies and the proportion of PDOs have a similar property in term of relationship with the weather. The collision frequencies and proportion of PDOs are higher in the winter time than in the summer time. However, this relationship is opposite for injury and fatal collision frequencies and proportions. During the summer time in Edmonton, drivers tend to speed more when the road condition is good and traffic is not congested.

4.3. *Impact of OPS24 on collisions*

In this section, we present the general impact of OPS24 on collisions by performing trend, intercept and weather factor analysis using interrupted time series with segmented regression model. In 2008, the OPS24 were held during the winter time (September, October, and December). However, due to the weather data availability, we started the analysis from October 2008.

Two examples of analysis are discussed here. For OPS24 on October 21, 2008, Fig. 4 (left) and Table 7 showed that there was no significant change in the level of injury and fatal collision proportions after OPS24, although the trend became significantly steeper.

Table 6
Parameters of Gaussian Models.

Explanatory variables	Injury and fatal collisions proportion		PDO proportion	
	Coeff	Pr(> | t |)	Coeff	Pr(> | t |)
Intercept	1.529e−01	<2e−16*	8.473e−01	<2e−16*
Slope	1.625e−03	0.6911	−7.777e−04	0.8489
Cosinus	7.635e−03	0.0494−	−8.535e−03	0.0279#
Sinus	−2.843e−03	0.3055	3.286e−03	0.2361
Temperature	1.647e−03	1.62e−11*	−1.623e−03	3.13e−11*
Snow	−5.751e−05	0.7478	2.966e−05	0.8682
Gust	−1.360e−04	0.3111	1.265e−04	0.3461
Null deviance	3.7213	0.7478	3.7317	
Residual deviance	2.9278		2.9317	
AIC	−3864.5		−3867.1	

Signif. codes:
* <0.01.
− <0.05.
<0.10.

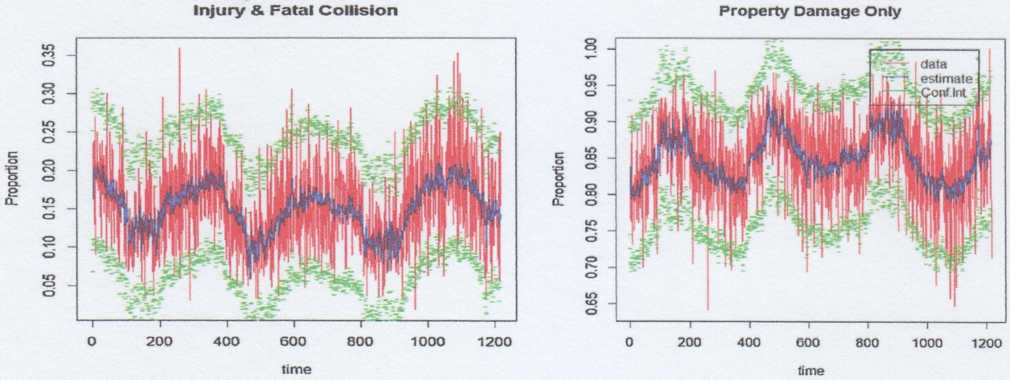

Fig. 3. Fitted model into the proportion for (left) injury and fatal proportion, (right) PDO proportion.

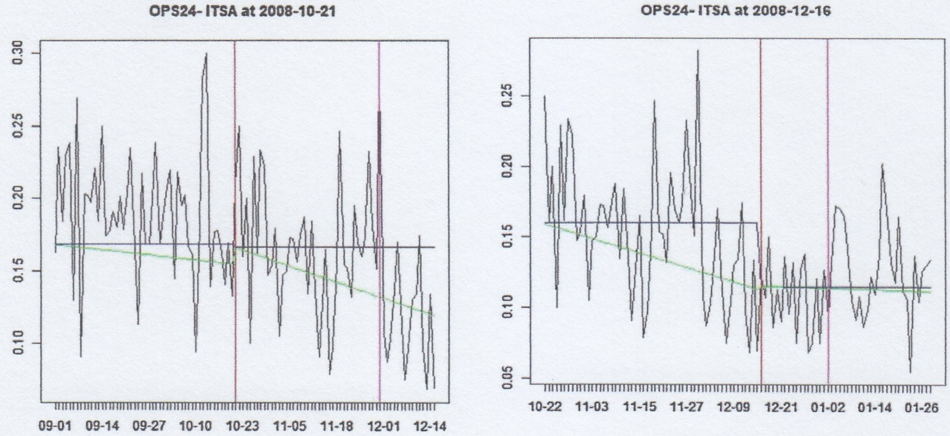

Fig. 4. Trend and intercept before/after OPS24 in 2008 (blue: intercept; green: trend; red: the OPS24 date; magenta: the change point). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 7
Trend and intercept analysis results 2008–2011.

Date	Change point	Mean before	Trend before	Mean after	Trend after	Temperature	Snow
10/21/2008	40	0.16842*	−0.00027	−0.00176	−0.00085^{+}	−0.00215	0.00323^{+}
12/16/2008	17	0.15978*	−0.00085*	−0.04473^{+}	−0.00009	−0.00295*	0.00384*
1/29/2009	25	0.12336*	−0.00004	0.00303	−0.00022	−0.00192	0.00292
3/10/2009	8	0.11768*	−0.00022	−0.01913	0.00176	−0.00364	0.00492
4/7/2009	20	0.11185*	0.00162	0.03539	−0.00080	−0.00201	0.00299
5/12/2009	4	0.12836*	−0.00074	−0.00339	−0.00044	−0.00313	0.00541^{+}
6/9/2009	9	0.11946*	−0.00015	−0.02880	0.00317	−0.00466	0.00546
6/26/2009	7	0.14802*	0.00340	0.01317	0.00103*	0.00291	−0.00170
9/29/2009	13	0.23054*	−0.00024	−0.04045	−0.00459^{+}	0.00401	−0.00329
10/16/2009	35	0.14672*	−0.00212	−0.00688	−0.00040	0.00067	0.00060
12/15/2009	27	0.13867*	−0.00051	−0.00881	−0.00024	−0.00013	0.00141
1/20/2010	13	0.11886*	−0.00010	−0.04137	0.00486	−0.00171	0.00308
2/5/2010	17	0.06027^{+}	0.00634*	0.06988*	0.00056	0.00217	−0.00248
3/29/2010	1	0.12452*	0.00067	0.02840	0.00345	0.00196	−0.00272
4/13/2010	6	0.13046*	0.00282	0.01170	−0.00008	−0.00189	0.00207
6/28/2010	40	0.12986*	0.00034	0.01254	−0.00036	−0.00076	0.00134
9/13/2010	42	0.14103*	0.00053	0.02260	−0.00006	0.00152	−0.00119
11/15/2010	26	0.17917*	−0.00017	−0.05020^{+}	0.00001	−0.00337*	0.00256
12/14/2010	28	0.11547*	0.00013	0.01362	−0.00020	−0.00186	0.00373*
1/19/2011	12	0.11140*	0.00002	−0.02274	0.00203	−0.00238	0.00369^{+}
2/4/2011	39	0.09215*	0.00181	0.01366	0.00060	−0.00157	0.00231
3/28/2011	21	0.12205*	0.00041	0.05790	−0.00110	−0.00041	0.00140
4/21/2011	14	0.18725*	0.00144	−0.05663	0.00135	0.00008	0.00184
5/19/2011	13	0.21358*	0.00068	0.03262	−0.00254	0.00752^{+}	−0.00625
6/7/2011	45	0.22788*	0.00504^{+}	−0.03288	0.00008	−0.00239	0.00094
11/15/2011	5	0.13324*	−0.00015	−0.03703	0.00226*	−0.00405*	0.00632*
12/16/2011	13	0.10786*	0.00214^{+}	0.02011	0.00177	−0.00266	0.00458

* Significant at level 5%.
$^{+}$ Significant at level 10%.

Table 8
Summary of collision trend and weather influence analysis results.

	Daily collisions	Daily injury and fatal collisions	Daily injury and fatal collision proportions	Daily PDO collisions	Daily PDO collision proportions
Linear trend	–	–	+	–	–
Periodic trend	Cosine	Cosine	Cosine	Cosine	Cosine
Mean Temperature	–	–	+	–	–
Thichness of snow	+	+	–	+	+
Wind gust	+	+	–	+	+

+, statistically significant positive association between dependent and independent variable.–, statistically significant negative association between dependent and independent variable, +/–, non-statistically significant positive/negative association between dependent and independent variable.

Snow on the ground showed significant influence on the pattern change. The mean of collisions proportions lasted up to 40 days after the OPS24 ended. The analysis of OPS24 on December 16, 2008 showed the opposite result: Fig. 4 (right) and Table 7 showed that the level of injury and fatal collision proportions after OPS24 changed significantly but not the trend. Both mean temperature and snow on the ground significantly influenced the fatal and injury collision proportions. The mean of collisions proportions lasted up to 17 days after the OPS24 ended.

Overall, OPS24 events as well as temperature and snow on the ground did not influence the injury and fatal collision proportions. This could happen since most of the OPS24s were held when snow was present, in which drivers tend to slow down. The schedules of OPS24 events sometimes were close to each other. These schedules were determined by the Edmonton Police Service based on several factors, including resources availability.

5. Conclusions

The negative binomial time series model that accounts for seasonality and weather factors showed that the mean temperature is negatively associated (with statistical significance) with all collision types, i.e., the higher the mean temperature, the lower the number of daily collisions. Thickness of snow is positively associated with injury and fatal collisions (with statistical significance), but is not significantly associated with PDO collisions. Wind gust is not associated with either injury and fatal collisions or PDO collisions. Furthermore, daily collisions, injury and fatal collisions showed a decreasing trend and the periodic trend followed the cosine function, which means that the collisions were high in January (winter), slowly decreased to June (summer), and then started to increase again in September (autumn).The results from this analysis were mainly similar to El-Basyouny and Kwon (2012)'s paper, although one of results was completely opposite. The similarities occurred in the following terms: in both studies, the injury and fatal collisions had a significant declining annual trend, significant inverse relationship with mean temperature, and significant positive relationship with thickness of snow. Also, in both studies, PDO collisions had significant inverse relationship with mean temperature and positive relationship with thickness of snow. The contradiction occurred in annual trend of PDO collisions: in El-Basyouny and Kwon (2012)'s paper, PDO collisions had significant growing annual trend, while in this study the PDO collisions had a significant declining annual trend. One possible reason is that El-Basyouny and Kwon used 2000–2010 data, where the trends were not uniform over this course of 11 years that could be due to several reasons, including changes in population, roadway structure, and traffic enforcement strategies

The Gaussian time series model that accounts for seasonality and weather factors showed that only mean temperature had a significant positive association with the daily injury and fatal collision proportions. The linear trend is not statistically significant for the daily injury and fatal collisions or PDO collision proportions.

Table 8 summarizes the association between collision and weather factors. El-Basyouny and Kwon (2012) did not analyze the relationship between the weather factors and the daily injury and fatal collision proportions.

The interrupted time series model with segmented regression model showed that there was no statistical evidence that OPS24 had impact on collision reduction. The change in collision pattern after OPS24 was due to the change in weather factors.

Given the fact that the City of Edmonton is characterized by its sometimes extreme winter weather which typically constitutes five to six month of the year (El-Basyouny and Kwon, 2012) and the magnitude of the problem caused by the adverse weather conditions, it is imperative to have a solid understanding of the association between collision and weather factors to come up with an optimum proactive traffic safety countermeasures. Detail analysis of various relevant countermeasures is beyond the scope of this study. However, a list of relevant countermeasures and their corresponding collision modification factors (CMFs) based on previous studies, such as installing icy curve warning system (CMF = 0.82 for total collisions) and increasing pavement friction (CMF = 0.76 for total collisions) can be reviewed from Crash Modification Factors Clearinghouse (Clearing House, 2013).

Acknowledgements

The authors are very grateful to Dr. Stevanus A. Tjandar, Dr. Yongsheng Chen and two anonymous referees for their valuable comments and suggestions, which contributed to improve the quality of the paper.

References

Andreescu, M., Frost, D.B., 1998. Weather, traffic accidents in Montreal, Canada. Climate Research 9, 225–230.
Andrey, J., Mills, B., Vandermolen, J.,2001. Weather Information and Road Safety: Final Report. Final Report Presented to the Institute for Catastrophic Loss Reduction.
Andrey, J., Yagar, S., 1993. A temporal analysis of rain related crash risk. Crash Analysis and Prevention 25, 465–472.
Box, G., Jenkins, G., 1970. Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco.
Brännäs, K., Johansson, P., 1994. Time series count data regression. Communications in Statistics A 23, 2907–2925.
Brodsky, H., Hakkert, A.S., 1988. Risk of a road accident in rainy weather. Crash Analysis and Prevention 20, 161–176.
Brown, B., Baass, K., 1997. Seasonal variation in frequencies and rates of highway accidents as a function of severity. Transportation Research Record 1581, 59–65.
Cameron, A.C., Trivedi, P.K., 1998. Regression Analysis of Count Data. Cambridge University Press.
Castro, M., Paleti, R., Bhat, C.R., 2012. A latent variable representation of count data models to accommodate spatial and temporal dependence: application to predicting crash frequency at intersection. Transportation Research Part B: Methodological 46 (1), 253–272.
Clearing House, Modification Factors Clearinghouse, http://www.cmfclearinghouse.org/index.cfm (accessed 4.4 13).
CRAN, The R-Project for Statistical Computing: Comprehensive R Archive Network, www.r-project.org, 2012.
Davis, R.A., Dunsmuir, W.T.M., Wang, Y., 2000. On autocorrelation in a poisson regression model. Biometrika 87, 491–505.

Davis, R.A., Wu, R., 2009. Negative binomial models for time series of counts. Biometrika 96 (3), 735–749.
Eisenberg, D., 2004. The mixed effects of precipitation on traffic crashes. Accident Analysis and Prevention 36, 637–647.
El-Basyouny, K., Kwon, D., 2012. Assessing Time and Weather Effects on Collision Frequency by Severity in Edmonton Using Multivariate Safety Performance Functions. In: Presented at the 91st Annual Meeting of Transportation Research Board.
Flores, B.E., 1986. A pragmatic view of accuracy measurement in forecasting. Omega 14 (2), 93–98.
Fridstrom, L., Liver, J., Ingebrigtsen, S., Kulmala, R., Thomsen, L., 1995. Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts. Crash Analysis and Prevention 27, 1–20.
Fridstrom, L., Ingebrigtsen, S., 1991. An aggregate accident model based on pooled, regional time-series data. Accident Analysis and Prevention 23, 363–378.
Halim, Jiang, S., Speed, H., 2012. Collision Impact Analysis of Operation 24 Hours in the City of Edmonton. City of Edmonton Office of Traffic Safety.
Hamed, M.M., Al-Eideh, B.M., Al-Sharif, M.M., 1999. Traffic accidents under the effect of the gulf crisis. Safety Science 33, 59–68.
Hawkins, D.M., Qiu, P., Kang, C.W., 2003. The changepoint model for statistical process control. Journal of Quality Technology 35, 335–365.
Hawkins, D.M., Zamba, K.D., 2005. Statistical Process Control for Shifts in Mean or Variance Using a Changepoint Formulation 47 (2), 164–173.
Hawkins, D.M., Deng, Q., 2010. A nonparametric change point control chart. Journal of Quality Technology 42 (2), 165–173.
Killick, R., Eckley, I.A., 2011. Changepoint: An R Package for Changepoint Analysis. Lancaster University.
Knapp, K., Smithson, L.D., Khattak, A.J., 2000. The mobility and safety impacts of winter storm events in a freeway environment. Mid-Continent Transportation Research Symposium, 67–71.
Lombardo L. Overview of US Crashes and Environment. Paper presented at Second Weather Information for Surface Transportation (WIST) Forum; December 4–6, 2000; Rockville, Md.
Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transportation Research Part A, http://dx.doi.org/10 1016/j.tra.2010.02. 001.
Miaou, S.P., Mallick, J.J., 2003. Roadway traffic crash mapping: a space time modeling approach. Journal of Transportation and Statistics 6 (1), 33–57.
National Research Council, 2004. Where the Weather Meets the Road? National Academies Press, Washington, DC.
Perrin, N.A., 2009. Analysis of Interrupted Time Series with Segmented Regression. Center of Health Research.
Perry, A.H., Symons, L.J., 1991. Highway Meteorology. E & F N Spon Press, Swansea, Wales.
Quddus, M.A., 2008. Time series count data models: an empirical application to traffic accidents. Accident Analysis & Prevention 40 (5), 1732–1741.
Usman, T., Fu., L., Miranda-Moreno, L.F., 2011. Accident prediction models for winter road safety: does temporal aggregation of data matter? Transportation Research Record: Journal of the Transportation Research Board 2237, 144–151.
Vanlaar, W., Robertson, R., Marcoux, K., 2008. The road safety monitor 2007: excessive speeding. Traffic Injury Research Foundation.
Vanlaar, W., Robertson, R., Marcoux, K., 2011. Evaluation of the Photo Enforcement Safety Program of The City of Winnipeg: Final Report. Traffic Injury Research Foundation.
Weather and Highways, 2004. Report of a Policy Forum. American Meteorological Society, Washington, DC.